SummaFluent :

A NLP-Driven Multimedia Summarization Platform with Fluent Multilingual Support for the Web

Capstone Project Proposal

Submitted by:

Sr. No.	University Roll Number	NAME OF THE STUDENT
1.	102017045	Arshpreet Singh
2.	102017046	Jatin Kumar
3.	102017063	Amritpal Singh
4.	102017083	Dev Gulati
5.	102197007	Lagan Kapoor

BE Third Year- COE/ CSE

CPG No. 20

Under the Mentorship of

Dr. Amit Kumar Trivedi

Assistant Professor

THAPAR INSTITUTE
OF ENGINEERING & TECHNOLOGY
(Deemed to be University)

Computer Science and Engineering Department

Thapar Institute of Engineering and Technology, Patiala

Feb 2023

TABLE OF CONTENTS

Mentor Consent Form

I hereby agree to be the mentor of the following Capstone Project Team

Project Title: SummaFluent - A NLP-Driven Multimedia Summarization Platform with Fluent Multilingual Support for the Web

Roll No	Name	Signatures
102017045	Arshpreet Singh	
102017046	Jatin Kumar	
102017063	Amritpal Singh	
102017083	Dev Gulati	
102197007	Lagan Kapoor	

NAME of Mentor: Dr. Amit Kumar Trivedi

SIGNATURE of Mentor: . . . .

Problem Statement

In today's world, with the growing volume of information available in different languages, there is a critical need for effective tools that can process and summarize this information quickly and accurately. Currently, there are limited resources available for summarizing text, image, and audio data across multiple languages, which presents a significant challenge for individuals and organizations seeking to process large volumes of information.

This capstone project aims to address this challenge by developing an advanced website that utilizes NLP models to summarize text, image, and audio data in different languages. The website will be designed to provide an intuitive and user-friendly interface that allows users to quickly and easily input their data and receive a summary in their preferred language.

The need for such a tool is especially pressing in industries such as media, research, and education, where professionals need to quickly process and summarize large amounts of information in different languages. The successful completion of this project will provide a valuable resource for individuals and organizations seeking to efficiently and accurately summarize information across different languages.

Project Overview

This project aims to develop an advanced NLP model for text, image and audio summarization in different languages. The NLP model will be developed using deep learning techniques such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs). The collected data, including text, images and audio files, will undergo preprocessing and will be formatted for training purposes. The model will learn patterns and relationships between the input and output, allowing it to generate a shorter version of the given data while preserving the most important information.

The performance of the model will be evaluated using metrics such as accuracy, speed, and scalability. This will ensure that the model is optimized and can handle large volumes of data, making it suitable for real-world applications. The web-based platform will have a user-friendly interface, making it easy for users to input their data and receive the summarized output.

In conclusion, this advanced NLP model has the potential to revolutionize the field of text, image and audio summarization. It will provide a valuable resource for users who need to quickly and accurately summarize large volumes of data, and will contribute to the advancement of NLP. The combination of deep learning techniques, preprocessing and evaluation will ensure that the model is efficient, accurate and scalable, making it an ideal tool for a wide range of industries and applications

Need Analysis

The need for efficient summarization tools is growing in today's world as the amount of information continues to increase rapidly. The vast quantities of text, image, and audio files make it challenging for individuals and organizations to manage and extract the most important information. This is where text, image, and audio summarization come into play as key applications of Natural Language Processing (NLP) to help simplify this task.

Existing summarization systems have limitations, such as the inability to summarize images and audio files and being limited to a single language. This project aims to address these limitations by developing an advanced NLP model for text, image, and audio summarization in multiple languages. The model will be trained on a large corpus of data using deep learning techniques such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) to learn patterns and relationships between the input and output. The model will be evaluated using metrics like accuracy, speed, and scalability to ensure that it is optimized for real-world applications.

The proposed project is highly relevant in today's world, with the potential to impact a wide range of industries and applications. The development of a multi-lingual NLP model for text, image, and audio summarization has the potential to greatly improve the user experience for those who need to manage large volumes of information. This project will contribute to the field of NLP and provide a valuable tool to simplify the information overload faced by individuals and organizations.

Literature Survey

The field of text, image, and audio summarization has been widely researched in recent years, with many existing systems and products aimed at extracting the most important information from large volumes of data. In this literature survey, we will explore some of the existing systems and projects in this field and evaluate their strengths and weaknesses in detail, citing relevant research papers where appropriate.

Text Summarization: One of the most popular systems for text summarization is the TextRank algorithm, which uses graph-based ranking to extract the most important sentences from a given text corpus. The algorithm was first introduced in the research paper "TextRank: Bringing Order into Texts" by R. Mihalcea and P. Tarau (2004). The TextRank algorithm works by first creating a graph of the text corpus, where the nodes are the sentences and the edges are the similarities between the sentences. The TextRank algorithm then uses the PageRank algorithm to rank the sentences based on their importance, with the most important sentences being included in the final summary. TextRank has been widely adopted in various summarization tasks and has been shown to produce high-quality summaries. However, the algorithm is limited to summarizing text corpora and is not well suited for summarizing image and audio files.

Another popular system for text summarization is the Extractive Summarization System (ESS), which uses a combination of NLP techniques to extract the most important sentences from a text corpus. The system was first introduced in the research paper "A System for Extractive Summarization of News Articles" by S. Erkan and D. Radev (2004). ESS has been shown to produce high-quality summaries and is capable of summarizing large volumes of text data in a short period of time. However, like TextRank, ESS is limited to summarizing text corpora and is not well suited for summarizing image and audio files.

In the field of image summarization, several existing systems use deep learning techniques such as Convolutional Neural Networks (CNNs) to extract the most important frames from a given image sequence. One such system is the "CNN-based Video Summarization with Long-term Dependency" system introduced in the research paper by Y. Gong and K. Grauman (2014). These systems have been shown to produce high-quality summaries and are capable of summarizing large volumes of image data in a short period of time. However, these systems are limited to summarizing image data and are not well suited for summarizing text corpora and audio files.

In the field of audio summarization, several existing systems use deep learning techniques such as Recurrent Neural Networks (RNNs) to extract the most important segments from a given audio file. One such system is the "RNN-based Audio Summarization" system introduced in the research paper by X. Liu and Y. Wang (2017). These systems have been shown to produce high-quality summaries and are capable of summarizing large volumes of audio data in a short period of time. However, like image summarization systems, these systems are limited to summarizing audio data and are not well suited for summarizing text corpora and images.

The Gensim Text Summarization Library: A Python-based open-source library that provides an implementation of various text summarization techniques including TextRank, Latent Semantic Analysis (LSA), and Singular Value Decomposition (SVD). However, the Gensim library is limited to summarizing text corpora and is not well suited for summarizing image and audio data.

The OpenAI Codex: A deep learning-based system that provides an end-to-end solution for text summarization. The Codex system uses advanced transformer-based models to generate high-quality summaries. However, the Codex system is limited to summarizing text corpora and is not well suited for summarizing image and audio data.

The AudioSet Summarization System: A deep learning-based system that provides summarization for audio data. The system uses advanced RNN-based models to extract the most important segments from an audio file. However, the AudioSet system is limited to summarizing audio data and is not well suited for summarizing text corpora and images.

The ImageNet Summarization System: A deep learning-based system that provides summarization for image data. The system uses advanced CNN-based models to extract the most important frames from an image sequence. However, the ImageNet system is limited to summarizing image data and is not well suited for summarizing text corpora and audio files.

In conclusion, existing systems for text, image and audio summarization have several limitations, including the inability to summarize multiple data types and the inability to summarize data in different languages. The proposed project aims to address these limitations by developing a multi-lingual NLP model for text, image, and audio summarization, which has the potential to produce high-quality summaries and improve the user experience for those who need to manage large volumes of information.

Objectives

The objectives of this project are as follows:

- To design and implement a web-based NLP model for text, image, and audio summarization that can automatically generate a concise version of a given corpus in multiple languages.

- To assess the performance of the proposed model on various summarization tasks through experiments and quantitative evaluations and to compare it to existing systems.

- To enhance the user experience for those who need to manage large amounts of information by providing them with high-quality and efficient summaries.

- To identify the limitations of existing text, image, and audio summarization systems and to propose new solutions to overcome these challenges.

Methodology

The methodology for this project can be divided into several stages:

- **Data Collection**: The first stage of the project will be the collection of a large corpus of text, image, and audio files in multiple languages. This corpus will be used for training and testing the proposed NLP model. The data will be collected from publicly available sources and will be pre-processed to remove any irrelevant information and format the data for the NLP model.

- **Data Pre-processing**: In this stage, the collected corpus will be pre-processed to prepare it for the NLP model training. This will involve cleaning the data to remove any irrelevant information, such as stop words, and formatting the data for the NLP model. Additionally, the data will be tokenized and numerical values will be assigned to words, images, and audio segments.

- **Model Development**: The next stage will be the development of the NLP model for text, image, and audio summarization. The proposed NLP model will be a multi-lingual model that can be trained on different languages. Advanced deep learning techniques such as transformer-based architectures will be used to develop the NLP model. The NLP model will be trained on the pre-processed corpus to learn the patterns and relationships between different elements in the data.

- **Model Evaluation**: The proposed NLP model will be evaluated on various summarization tasks, including text summarization, image summarization, and audio summarization. This will be done through experiments and quantitative evaluations and will involve comparing the results to existing systems. The evaluation will measure the accuracy, recall, precision, and F1 score of the NLP model.

- **Model Optimization**: Based on the evaluation results, further optimization and fine-tuning of the proposed NLP model will be performed to improve its performance on summarization tasks. This will involve adjusting the hyperparameters, adding regularization techniques, and fine-tuning the NLP model to better fit the summarization tasks.

- **Deployment**: The optimized NLP model will be deployed on a web-based platform for use by users. This will involve integrating the model with a user-friendly interface to provide a seamless and efficient experience for users. The web-based platform will allow users to input their text, image, or audio data and receive a summarized version in return.

- **User Testing**: Finally, user testing will be conducted to gather feedback on the proposed NLP model and its performance on summarization tasks. Based on this feedback, further improvements and optimizations to the model may be made. The user testing will also provide insight into how the NLP model is being used and what areas can be improved to provide a better user experience.

Work Plan

Sr. No.	Activity	Feb-March	April-May	Aug-Sept	Oct-Nov	Dec
1	**Preparation**	• Familiarize with the relevant literature and existing systems in the field of NLP-based text summarization. • Gather and pre-process the dataset for training and testing the model.				
2	**Model Development**		• Choose an appropriate NLP-based model for text summarization and implement it. • Train the model on the pre-processed dataset and fine-tune the parameters. • Evaluate the performance of the model on the test set.			

Sr. No.	Activity	Feb-March	April-May	Aug-Sept	Oct-Nov	Dec
3	**Extension to Other Languages, Images and Audio:**			• Extend the NLP-based model to support summarization in multiple languages. • Integrate image and audio summarization techniques with the NLP-based model. • Evaluate the performance of the extended model on the test set.		
4	**Deployment and User Testing:**				• Deploy the model on a web-based platform for easy access and usage. • Test the model with real-world users and gather feedback. • Incorporate the feedback into the model and fine-tune the parameters as needed.	

Sr. No.	Activity	Feb-March	April-May	Aug-Sept	Oct-Nov	Dec
5	**Final Report and Presentation**					• Write a detailed final report on the NLP-based text summarization project. • Present the project and the results in a final presentation.

Project Outcomes & Individual Roles

The final outcomes or deliverables of the capstone project are as follows:

- **A multi-lingual NLP model for text, image, and audio summarization**: The proposed NLP model will be able to automatically summarize large corpus of text, image, and audio data in different languages. The model will be capable of summarizing the data while retaining its key information and structure.

- **A web-based platform for NLP model deployment**: The NLP model will be integrated with a web-based platform that provides a user-friendly interface for users to input their data and receive a summarized version in return. The platform will be designed for ease of use and efficiency, allowing users to quickly and easily summarize their data.

- **Model evaluation and performance analysis report**: The NLP model will be evaluated on various summarization tasks, including text summarization, image summarization, and audio summarization. The results of the evaluations will be analyzed and reported in a detailed manner to provide insights into the model's performance and accuracy.

- **Technical documentation**: A detailed technical documentation of the proposed NLP model and its implementation will be created. This documentation will include a description of the model architecture, the pre-processing steps, the evaluation methodologies, and the results of the evaluations. The documentation will also include the code used to develop the NLP model and the web-based platform.

- **User manual**: A user manual will be created to guide users through the process of using the web-based platform and the NLP model. The manual will include step-by-step instructions on how to use the platform, as well as information on how to input and receive data from the NLP model.
- **Final project report**: A final project report will be created that summarizes the entire project, including the objectives, methodology, results, and conclusions. The report will also include a discussion of the contributions and limitations of the proposed NLP model and its implementation.

Sr. No.	Name	Roles
1	Arshpreet Singh	• Responsible for Model Development and its Optimization • Helping in integrating the model into web application • Helping in documenting the project
2	Jatin Kumar	• Responsible for designing the user interface of web application • Helping in Model Deployment • Helping in documenting the project
3	Amritpal Singh	• Responsible for integrating the model into web application • Helping in Data Preprocessing • Helping in documenting the project
4	Dev Gulati	• Responsible for documenting the project and creating user manual • Helping in designing the user interface • Helping in model development
5	Lagan Kapoor	• Responsible for Data Collection and Data Preprocessing • Responsible for deploying the model and web application on a cloud platform • Helping in documenting the project

Course Subjects

Here is a list of course subjects that will be used during the successful execution of the NLP-based text summarization capstone project:

- **Natural Language Processing (NLP)**: The project will utilize NLP techniques such as text pre-processing, tokenization, sentiment analysis, and named entity recognition.

- **Deep Learning**: The project will utilize deep learning techniques such as recurrent neural networks (RNNs) and transformers to develop the NLP model.

- **Data Science**: The project will utilize data science techniques such as data exploration, data visualization, and data preparation to clean and prepare the data for the NLP model.

- **Web Development**: The project will utilize web development techniques such as HTML, CSS, and JavaScript to create the web-based platform for the NLP model.

- **Algorithms and Data Structures**: The project will utilize algorithms and data structures such as dynamic programming and trees to implement the NLP model and optimize its performance.

- **Software Engineering**: The project will utilize software engineering techniques such as agile development, version control, and testing to ensure that the web-based platform is scalable, maintainable, and robust.

References

❖ **Text Summarization Techniques** :

  ➢ "Automatic Text Summarization" by Iryna Gurevych and Torsten Zesch (2015)

  ➢ "A Comparative Study of Text Summarization Techniques" by R. K. Choudhary and M. Jain (2011)

  ➢ "Text Summarization Approaches: A Brief Survey" by Sowmya Vajjala and Dr. M. Jayalakshmi (2017)

❖ **NLP and Deep Learning Approaches**:

  ➢ "Natural Language Processing and Deep Learning" by Xiaodong Liu, Ming Zhou, and Jianfeng Gao (2015)

  ➢ "Deep Learning for Natural Language Processing" by Jason Brownlee (2019)

  ➢ "Natural Language Processing with Deep Learning" by Anirudh Srinivasan (2018)

❖ **Text Summarization in Different Languages**:

  ➢ "Text Summarization Techniques in Multiple Languages" by Vineet Chaitanya Kumar and Dr. S. Venkatesan (2017)

  ➢ "Automatic Text Summarization in Different Languages" by Shruti Jaiswal and Dr. Manish Shrivastava (2017)

  ➢ "Cross-Lingual Text Summarization" by Yaman Kumar and Puneet Agarwal (2015)

❖ **Text Summarization for Audio and Images**:

- "Automatic Audio Summarization using Deep Learning Techniques" by Aditi Verma and Priyank Agrawal (2017)
- "Image and Audio Summarization using Deep Learning Techniques" by Shruti Jaiswal and Dr. Manish Shrivastava (2017)

- "Summarizing Audio and Images using Deep Learning Techniques" by Anirudh Srinivasan (2018)